**CENTRE DE RECHERCHE SUR LES TRANSPORTS**
(C.R.T.)
*CENTRE FOR RESEARCH ON TRANSPORTATION*

UNIVERSITÉ DE MONTRÉAL
ÉCOLE DES HAUTES ÉTUDES COMMERCIALES
ÉCOLE POLYTECHNIQUE

CENTRE DE RECHERCHE SUR LES TRANSPORTS
UNIVERSITÉ DE MONTRÉAL
C.P. 6128, SUCCURSALE A
MONTRÉAL
CANADA H3C 3J7

TÉLÉPHONE : (514) 343-7575
EMAIL : crt@crt.umontreal.ca
TÉLÉCOPIEUR (FAX) : (514) 343-7121

CRT-788

# The Integer L-Shaped Method for Stochastic Integer Programs with Complete Recourse

by

Gilbert Laporte[1]
François V. Louveaux[2]

[1] Centre de recherche sur les transports, Université de Montréal, C.P. 6128, succursale A, Montréal, Québec, Canada H3C 3J7

[2] Facultés Universitaires Notre-Dame de la Paix, Faculté des Sciences Économiques et Sociales, 8 Rempart de la Vierge, 5000 Namur, Belgium.

Centre de recherche sur les transports — Publication #788

June 1991
Revised January 1992 and November 1992

Dépôt légal – Bibliothèque nationale du Québec,
Bibliothèque nationale du Canada, 1992

93-01803/


## Abstract

In this paper, a general branch-and-cut procedure for stochastic integer programs with complete recourse and first stage binary variables is presented. It is shown to provide a finite exact algorithm for a number of stochastic integer programs, even in the presence of binary variables or continuous random variables in the second stage.


Key words : Stochastic integer programming, L-shaped method, branch-and-cut.


## Résumé

On décrit dans cet article une procédure générale de "branch-and-cut" pour la résolution de programmes stochastiques en nombres entiers avec recours complet et variables binaires en première période. La méthode proposée constitue un algorithme exact fini, applicable à une classe générale de programmes stochastiques en nombres entiers, même lorsque ceux-ci comportent des variables binaires ou des variables aléatoires continues en seconde période.


Mots-clés : Programmation stochastique en nombres entiers, méthode "L-shaped", "branch-and-cut".

## 1. Introduction

Stochastic integer programs have a well-deserved reputation for being computationally intractable. In fact, they combine two types of programs which by themselves are often difficult to solve. Stochastic programs are modeled in two stages. In the first stage, a decision has to be taken without full information on some random variable. When the value taken by the random variable is known, corrective or recourse actions can be taken in the second stage. Stochastic integer programs are those for which some variables are restricted to be integer, either in the first stage or in the second stage.

To the best of our knowledge, the first proposed solution method was a cutting planes algorithm by Wollmer (1980) for the case where second stage decision variables are continuous and the random variables are discrete. In a similar vein, Louveaux (1986) has shown that multistage programs having the block-separability property are equivalent to two-stage problems with first stage discrete variables only. When block-separability does not hold, it is still possible to construct an extended first stage problem by including the extensive form of the second stage discrete variables. An example of this procedure is provided by Bienstock and Shapiro (1988). It is only applicable for problems with a limited number of realizations of the random variables. Other authors have studied properties of the expected value of the second stage program when it involves integer decision variables. General results concerning the continuity of this function have been obtained by Stougie (1985). This type of research has also been done for specific applications: e.g. Jaillet (1988), or Laporte and Louveaux (1990) for the analysis of some vehicle routing problems. Several studies have addressed the behavior of various heuristics for specific problems (see, e.g., chapters 8 and 21 in Ermoliev and Wets (1988), also containing general results on stochastic programming). One line of research is the study of the asymptotic optimality of some heuristics, based on the probabilistic analysis

of algorithms. However, Psaraftis (1984) has shown through an example that optimality is truly asymptotic in the sense that it only holds for problems of extremely large scale. The need for a practical solution method of general applicability still exists.

The object of this paper is to provide such an algorithm for problems with first stage binary decision variables. The random vectors and the second stage decision variables may be discrete or continuous, and the application of the method varies accordingly.

The structure of the paper is as follows. In Section 2, we review the definition of two-stage stochastic linear programs with recourse. We also provide a simple taxonomy for the various types of stochastic integer programs. The general branch-and-cut procedure, called "Integer L-shaped" due to its similarity with the well-known methods of Van Slyke and Wets (1969) for continuous programs, or Benders' decomposition (1962), is presented in Section 3. We prove its finite convergence in the case of first stage binary decision variables. In Section 4, we discuss the implementation of the proposed method to problems with various types of random vectors and second stage decision variables. Improvements are proposed in Section 5 and an example is provided in Section 6.

## 2. Two-Stage Stochastic Linear Programs with Recourse

The general formulation of a two-stage stochastic linear program with fixed recourse is to find

$$\min z = cx + E_\xi \min \{q(\omega)y \mid Wy = h(\omega) - T(\omega)x, y \in Y\} \tag{1}$$

$$s.t. \quad Ax = b$$
$$x \in X,$$

where $c$ is a known vector in $\mathbb{R}^{n_1}$, $b$ a known vector in $\mathbb{R}^{m_1}$, $\xi$ a random $N$-vector defined on the probability space $(\Omega, \mathrm{A}, P)$, and $A$ and $W$ are known matrices of size $m_1 \times n_1$ and $m_2 \times n_2$, respectively. For each $\omega, T(\omega), q(\omega)$ and $h(\omega)$ have conformable

sizes. Piecing together the stochastic components of the problem, we obtain a vector $\xi(\omega) = (q(\omega), h(\omega), T_1(\omega), ..., T_{m_2}(\omega))$ with $N = n_2 + m_2 + (m_2 \times n_1)$ components, where $T_i(\omega)$ is the $i^{th}$ row of $T(\omega)$. Transposes have been eliminated for simplicity. $E_\xi$ represents the mathematical expectation with respect to $\xi$. Let also $\Xi \subseteq \mathbb{R}^N$ be the support of $\xi$, i.e. the smallest closed subset in $\mathbb{R}^N$ such that $P(\xi \in \Xi) = 1$. If $\xi$ has finite support, there exists a discrete probability distribution which describes $\xi$.

The above problem is equivalent to the so-called deterministic equivalent program:

$$\min \; z = cx + Q(x) \tag{2}$$

$$s.t. \qquad Ax = b, x \in X$$

where
$$Q(x) = E_\xi \, Q(x, \xi(\omega)) \tag{3}$$

and
$$Q(x, \xi(\omega)) = \min_y \{q(\omega)y \mid Wy = h(\omega) - T(\omega)x, y \in Y\}. \tag{4}$$

The function $Q(x)$ is called the second stage value function. The computation of $Q(x)$ when $Q(x, \xi)$ is unbounded or infeasible for some $\xi$ is similar to the continuous case. We also assume relatively complete recourse, i.e., $\{x \mid Ax = b, x \in X\} \subseteq \{x \mid Q(x) < \infty\}$ or more simply, complete recourse $\{Wy, y \in Y\} = \mathbb{R}^{m_2}$. Of particular interest in this paper are the cases when the sets $X$ or $Y$ define restrictions requiring some variables to be integer.

In order to characterize the various possible situations, we use the notation $a \mid b \mid c$ to represent a particular class of two-stage programs. The first letter characterizes the first stage variables, and is either $C$ if all first stage variables are continuous, $B$ if they are all binary, or $M$ when some first stage variables are continuous and others are binary (mixed integer problems). Similarly, the second letter characterizes the second stage variables with an identical representation. The third letter is either $D$ when $\xi$ has a finite support (discrete distribution), or $C$ when $\xi$ is a continuous random variable. Thus, $C \mid C \mid D$ represents a two-stage continuous stochastic program with a discrete distribution. For this case, well-

known characterization theorems and solution algorithms exist (see, e.g., Wets, 1983). For some classes of problems, it is also useful to know whether or not they possess the block-separable property (see Louveaux, 1986) and this is denoted by an additional $| s$. For example, $M \mid M \mid D \mid s$ is a stochastic program with mixed integer first and second stage variables, discrete random variables and block separable second stage. Note that here, no distinction is made between integer and binary variables, since integer variables can be transformed into binary variables.

## 3. The Integer L-Shaped Method

Let $\overline{X}$ be the polytope defined by the set of constraints in $X$, other than those defining the first stage variables type. Hence $X = \overline{X} \cap \mathbb{R}^{n_1}_+$ for $C \mid * \mid *$ problems, $X = \overline{X} \cap Z^{n_1}_+$ for $B \mid * \mid *$ problems, and similarly for $M \mid * \mid *$ problems. In this section, we assume that at least some first stage variables are binary. For simplicity, all variables are non-negative.

At a given stage of the method, we consider the so-called *current problem*:

$$\text{(CP)} \qquad \min \quad cx + \theta \tag{5}$$

$$s.t. \quad Ax = b \tag{6}$$

$$D_k x \geq d_k \qquad (k = 1, ..., s) \tag{7}$$

$$E_k x + \theta \geq e_k \qquad (k = 1, ..., t) \tag{8}$$

$$x \geq 0, \theta \in \mathbb{R}. \tag{9}$$

In this current problem, constraints (7) are used to give an appropriate representation of $\overline{X}$. Usually, $\overline{X}$ is relaxed at the initial stage, and additional constraints (7) are added when needed. These constraints are referred to as *feasibility cuts*. A set of feasibility cuts is said to be valid

if there exists some finite value $s$ such that $x \in \overline{X}$ if and only if $\{D_k x \geq d_k, k = 1, ..., s\}$. Examples of such relaxed constraints are induced constraints for second stage feasibility, or subtour elimination constraints in vehicle routing problems.

Variable $\theta$, along with the constraints (8), are used to express $\theta$ as an appropriate approximation of $Q(x)$. Constraints (8) are referred to as *optimality cuts*. A set of $t$ optimality cuts is said to be valid if for all $x \in X$, $(x, \theta) \in \{(x, \theta) \mid E_k x + \theta \geq e_k, k = 1, ..., t\}$ implies $\theta \geq Q(x)$. In practice, it is not required to have $\theta \geq Q(x)$ for all feasible $(x, \theta)$, but simply for the optimal $(x^*, \theta^*)$.

We first present the Integer L-shaped method in general terms.

Step 0 : Set $s := t := \nu := 0, \overline{z} := \infty$. The value of $\theta$ is set equal to $-\infty$ or to an appropriate lower bound and is ignored in the computation. The only pendant node corresponds to the initial current problem.

Step 1 : Select some pendant node in the list; if none exists, stop.

Step 2 : Set $\nu := \nu + 1$; solve the current problem. If the current problem has no feasible solution, fathom the current node; go to Step 1. Otherwise, let $(x^\nu, \theta^\nu)$ be an optimal solution.

Step 3 : Check for any relaxed constraint violation. If one exists, add one feasibility cut (7), set $s := s + 1$, and return to Step 2. If $cx^\nu + \theta^\nu > \overline{z}$, fathom the current problem and return to Step 1.

Step 4 : Check for integrality restrictions. If one is violated, create two new branches following the usual branch and cut procedure; append the new nodes to the list of pendant nodes; return to Step 1.

Step 5 : Compute $Q(x^\nu)$ and $z^\nu = cx^\nu + Q(x^\nu)$. If $z^\nu < \overline{z}$, then update $\overline{z}$.

Step 6 : If $\theta^\nu \geq Q(x^\nu)$, then fathom the current node and return to Step 1. Otherwise impose one optimality cut (8), set $t := t + 1$ and return to Step 2. ■

Bender's decomposition has long been applied to deterministic mixed integer programs or to stochastic programs with continuous variables. The proposed algorithm is an extension to the case of stochastic mixed integer programs. It differs from the classical L-shaped method since first-stage binary variables are now taken into account. Moreover, it can be seen that the fathoming rules are also different from what is usually done in branch-and-bound. In particular, nodes are not necessarily fathomed when integrality conditions are satisfied. The following proposition follows immediately from the fathoming rules and from the definitions of valid sets of cuts.

Proposition 1. For any problem for which a valid set of feasibility cuts and a valid set of optimality cuts exist, the Integer L-shaped method yields an optimal solution (when one exists) in a finite number of steps. ■

For most combinatorial problems of interest, valid feasibility cuts can be borrowed directly from the deterministic setting. As for optimality cuts, a valid set is known to exist for two-stage stochastic linear programs in continuous decision variables with discrete random variables. In the L-shaped method (Van Slyke and Wets, 1969), these are defined using the optimal simplex multipliers of the second stage recourse problem. It follows that the Integer L-shaped method is applicable to all problems belonging to the $* \mid C \mid D$ class. It offers an interesting alternative to the cutting planes method used by Wollmer (1980). Moreover, as a consequence of a Proposition by Louveaux (1986, Section 5), the method is valid for $* \mid M \mid D \mid s$ problems when the integer variables belong to the aggregated level decisions. Note that this property also applies to multistage programs. For several families of deterministic integer problems, branch-and-bound techniques are known to perform better than

cutting planes and we can therefore expect the same to hold for stochastic integer problems.

However, our purpose here is not to make further developments in that direction, but rather to derive optimality cuts for problems with first stage binary decision variables. Since these optimality cuts can be applied in cases where some second-stage variables are also binary, and in cases where the second-stage random variables are continuous, their use in the integer L-shaped method covers a wider class of applications than do existing methods. Two assumptions are required.

Assumption 1. Given a binary first stage vector of decision variables $x$, the function $Q(x)$ is computable from $x$. (This is frequently the case in location and routing problems, for example).

Assumption 2. There exists a finite value $L$ satisfying $L \leq \min_{x}\{Q(x) \mid Ax = b,\ x \in \{0,1\}^{n_1}\}$. (This is clearly the case if the original problem is bounded).

Since the first stage decision variables are binary, there are only a finite number of feasible first stage solutions. Let $r = 1, ..., R$ index these feasible solutions.

Proposition 2. Let $x_i = 1, i \in S_r$ and $x_i = 0, i \notin S_r$ be the $r^{th}$ feasible solution, and $\theta_r$ the corresponding expected second-stage value. Define the optimality cut as

$$\theta \geq (\theta_r - L)\left(\sum_{i \in S_r} x_i - \sum_{i \notin S_r} x_i\right) - (\theta_r - L)(|S_r| - 1) + L, \tag{10}$$

where $|S_r|$ is the cardinality of $S_r$. Then, the set of optimality cuts (10) defined for all first stage feasible solutions is a valid set of optimality cuts.

Proof. The quantity $\sum_{i \in S_r} x_i - \sum_{i \notin S_r} x_i$ is always less than or equal to $|S_r|$. It takes the value $|S_r|$ only when $x$ is the $r^{th}$ feasible solution. Now, when $\sum_{i \in S_r} x_i - \sum_{i \notin S_r} x_i$ is equal to $|S_r|$, the right-hand side takes the value $\theta_r$, and otherwise

the right-hand side takes a value less than or equal to $L$. A set of valid cuts is therefore obtained by imposing one such constraint for each first-stage feasible solution. ■

It follows from the proof of Proposition 2 that the procedure is finite as Proposition 1 applies.

## 4. Implementation

The applicability of the Integer L-shaped method is directly related to Assumptions 1 and 2 being satisfied. Also, as for all branch-and-cut methods, its efficiency is greatly improved if in addition to optimality cuts, valid constraints can be generated to provide a lower bounding functional on $Q(x)$. This is now discussed.

The validity of Assumption 1 is obviously problem dependent. It generally holds when the random vector has a discrete support of small cardinality, especially if only a limited number of scenarios are considered in the second stage. When the random vector is continuous, one can resort to approximation methods normally used in continuous stochastic optimization (Birge and Wets, 1986), such as discretizing continuous random variables. Moreover, there exists a wide variety of problems for which $Q(x)$ can be easily computed. Several interesting examples are found in the literature on vehicle routing problems (VRPs). These include the VRP with stochastic demands (Dror and Trudeau, 1986), the probabilistic traveling salesman problem (Jaillet, 1988), and the VRP with stochastic travel times (Laporte, Louveaux and Mercure, 1992).

We now discuss Assumption 2 and the derivation of a lower bounding functional on $Q(x)$. Let

$$Q(x,\xi) = \min_{y} \{qy \mid Wy = h - Tx, y \geq 0 \text{ and integer}\}. \tag{11}$$

As usual, $\xi = (q, h, T), Q(x, \xi)$ is set equal to infinity if no second stage solution exists, and to minus infinity if the problem is unbounded. Also define

$$R(x, \xi) = \min_{y} \{qy \mid Wy = h - Tx, y \geq 0\}, \tag{12}$$

the solution to the continuous relaxation of $Q(x, \xi)$, with similar conventions for the infeasible and unbounded cases. The expectations of these two functions are

$$Q(x) = E_{\xi} Q(x, \xi) \tag{13}$$

and

$$R(x) = E_{\xi} R(x, \xi), \tag{14}$$

respectively, with the classical convention "$(+\infty) + (-\infty) = (+\infty)$".

Proposition 3. Any continuous L-shaped optimality cut is a lower bound on $Q(x)$.

Proof. First note that $Q(x, \xi) \geq R(x, \xi)$ holds for all $x$ and $\xi$. Indeed, any feasible solution for (11) is feasible for (12), so that $R(x, \xi) = \infty$ only if $Q(x, \xi) = \infty$. If $Q(x, \xi)$ is finite, so is $R(x, \xi)$ and the result holds. Taking expectations on both sides yields $Q(x) \geq R(x)$. The conclusion follows by applying the continuous L-shaped optimality cut to $R(x)$, namely

$$Q(x) \geq R(x) \geq R\left(x^k\right) + \partial R\left(x^k\right)\left(x - x^k\right), \tag{15}$$

where $x^k$ is any feasible first stage solution, and $\partial R(x^k)$ is a subgradient of $R$ at $x^k$.

■

Corollary 4. A cut for $Q(x)$ is given by

$$\theta \geq R(x^k) + \partial R(x^k)(x - x^k). \tag{16}$$

■

Corollary 5. A lower bound on $Q(x)$ is given by

$$L = \min_{x,\theta} \{\theta \mid Ax = b, 0 \leq x \leq 1 \text{ and } (\theta, x) \text{ satisfies (16) for } k = 1, ..., s\}, \qquad (17)$$

for any finite value $s \geq 1$. ■

Note that (17) defines a finite lower bound on $Q(x)$ for any $s \geq 1$, as long as a feasible $x^k$ exists. Indeed, in such a case, $R(x^k)$ is finite and the last term of (16) is bounded below since $0 \leq x \leq 1$. Another observation is that (16) plays a different role from (10). An optimality cut ensures that $\theta \geq Q(x^k)$, but only at the particular point $x^k$ at which it has been defined. For any other integer point, it reduces to the weaker constraint $\theta \geq L$. On the other hand, the cut defined by (16) is more general in that it may provide a non-trivial lower bound on $Q(x)$, for several values $x$. However, when applied to the point $x^k$, at which it has been defined, it only guarantees that $\theta \geq R(x^k)$. In general, $R(x^k) < Q(x^k)$ since these two values coincide only if the solution to (11) and (12) at $x^k$ reveals no integrality gap for all values of $\xi$. As such, cuts (16) do not provide a valid finite set of optimality cuts.

Finally, as for the computation of $Q(x)$, there also exist several problems for which the lower bound $L$ and a lower bounding functional can easily be determined by arguments other than those used to obtain Corollaries 4 and 5. Applications of these principles and implementation details can be found in two recent studies on the VRP with stochastic travel times (Laporte, Louveaux and Mercure, 1992) and on a location problem with stochastic demands (Laporte, Louveaux and Van hamme, 1992).

## 5. Improved Optimality Cuts

The optimality cut defined by (10) can be improved by examining bounds on the value of the recourse function associated with neighbors of the current solution. As in Proposition 2, consider $S_r$, the set of variable indices for which $x_i = 1$ in the curent solution. Define the set

$N_r(s)$ of so-called $s$-neighbors as the set of solutions such that $\left(\sum_{i\in S_r} x_i - \sum_{i\notin S_r} x_i\right) = |S_r| - s$. Thus for $s = 1$, either for one $i \in S_r$ the variable $x_i$ becomes 0, or for one $i \notin S_r$ it becomes 1. Also define $\lambda_r(s)$ as a lower bound on $\min_{x\in N_r(s)} \{Q(x)\}$.

Proposition 6. Consider the $r^{th}$ feasible solution, $\theta_r$ the expected value of the second-stage function, and the set $N_r(1)$ of 1–neighbors. Also define $a = \max(\theta_r - \lambda_r(1), (\theta_r - L)/2)$. Then the following optimality cut holds:

$$\theta \geq a\left(\sum_{i\in S_r} x_i - \sum_{i\notin S_r} x_i\right) + \theta_r - a|S_r|. \tag{18}$$

Proof. Let $\delta = \left(\sum_{i\in S_r} x_i - \sum_{i\notin S_r} x_i\right)$. The right-hand side of (18) is equal to $\theta_r - as$ for an $s$-neighbor, i.e., when $\delta = |S_r| - s$. This is a valid lower bound on $Q(x)$. It is obvious for $s = 0$. When $s = 1, \theta_r - a$ is by construction bounded above by $\theta_r - (\theta_r - \lambda_r(1)) = \lambda_r(1)$. When $s = 2, \theta_r - 2a$ is bounded above by $\theta_r - 2(\theta_r - L)/2 = L$. Now, $\theta_r - sa \leq L$ for $s \geq 3$ since $a \geq 0$. ■

Convergence is again guaranteed by $\theta \geq \theta_r$ if $\delta = |S_r|$, and (18) improves on (10) for all 1–neighbors. Moreover, a further improvement for $s$-neighbors ($s > 1$) is sometimes possible.

Proposition 7. Consider the $r^{th}$ feasible solution, $\theta_r$ the expected value of the second-stage function, and $\lambda_r(s)$. If

$$(\theta_r - \lambda_r(s)) \leq s(\theta_r - \lambda_r(1)) \tag{19}$$

holds for all $s \leq (\theta_r - L)/(\theta_r - \lambda_r(1))$, with $\lambda_r(1) < \theta_r$, then (18) holds with $a = \theta_r - \lambda_r(1)$.

Proof. For an $s$-neighbor, (18) becomes $\theta \geq a(|S_r| - s) + \theta_r - a|S_r| = \theta_r - as$. The right-hand side is a valid lower bound on $Q(x)$ where $x \in N_r(s)$ since (19) implies

$\theta_r - as \leq \theta_r - (\theta_r - \lambda_r(s)) = \lambda_r(s)$ for $a = \theta_r - \lambda_r(1)$ and $s \leq (\theta_r - L)/(\theta_r - \lambda(1))$.

When $s > (\theta_r - L)/(\theta_r - \lambda_r(1))$, then $\theta_r - as \leq L$ and (18) remains valid. ■

## 6. Example

The implementation of (18) rests on the computation of $\lambda_r(1)$ (and possibly of $\lambda_r(s), s \geq 2$) and is therefore problem dependent. It is now illustrated on the following problem: let $i = 1,...,n$ denote $n$ inputs, and $j = 1,...,m$ denote $m$ outputs. Each input can be used to produce any output. First-stage decisions are represented by binary variables $x_{ij}$ with costs $c_{ij}$, and equal to 1 if $j$ is produced by $i$, and equal to 0 otherwise. Other first stage variables are binary $z_i$, with costs $f_i$, equal to 1 if and only if input $i$ is used for at least one output. The level of output $j$ obtained when $x_{ij} = 1$ is a random variable $\xi_{ij}$. A penalty $p_j$ is incurred whenever the level of output $j$ falls below a required treshold $d_j$; in such a case the binary second-stage variable $y_j^\xi$ takes the value 1. The problem is then a $B \mid B \mid *$ program of the form:

$$\min \sum_{i=1}^{m} f_i z_i + \sum_{i=1}^{n} \sum_{j=1}^{m} c_{ij} x_{ij} + E_\xi \left( \min \sum_{j=1}^{m} p_j y_j^\xi \right) \tag{20}$$

$$s.t. \qquad \sum_{i=1}^{n} \xi_{ij} x_{ij} + d_j y_j^\xi \geq d_j \qquad (j = 1,...,m) \tag{21}$$

$$x_{ij} \leq z_i \qquad (i = 1,...n; j = 1,...,m) \tag{22}$$

$$x_{ij}, y_j^\xi \in \{0,1\} \qquad (i = 1,...,n; j = 1,...,m; \xi \in \Xi). \tag{23}$$

In this problem,

$$Q(x) = \sum_{j=1}^{m} p_j P \left( \sum_{i \in I(j)} \xi_{ij} < d_j \right), \tag{24}$$

where $I(j) = \{i : x_{ij} = 1\}$. (Similar closed form expressions for $Q(x)$ are often encountered in problems with binary first-stage variables such as vehicle routing problems, for example). A first possibility is to use Proposition 7 for $Q(x)$. Define the set $N_r^+(1)$ as the subset of 1–neighbors for which all current positive $x_{ij}$'s remain unchanged and one $x_{ij}$ currently equal to 0 becomes equal to 1. Observe that among all 1–neighbors of the current solution, only those in $N_r^+(1)$ can reduce $Q(x)$. From (24),

$$\theta_r - \lambda_r(1) \leq \max_j \left\{ p_j P\left( \sum_{i \in I(j)} \xi_{ij} < d_j \right) \right\}. \tag{25}$$

Also observe that in this example, condition (19) applies. Indeed, $\theta_r - \lambda_r(s)$ is determined as the sum of the $s$ largest values of $\left\{ p_j P\left( \sum_{i \in I(j)} \xi_{ij} < d_j \right) \right\}$. It follows that $\lambda_r(s-1) - \lambda_r(s)$ is non-increasing in $s$, which is sufficient for (19) to hold. A second possibility is to use the separability of $Q(x)$ in (24) as

$$Q(x) = \sum_{j=1}^{m} Q_j(x), \tag{26}$$

with

$$Q_j(x) = p_j P\left( \sum_{i \in I(j)} \xi_{ij} < d_j \right). \tag{27}$$

Using a multicut approach as in Birge and Louveaux (1988), we define

$$\theta = \sum_{j=1}^{m} \theta_j \tag{28}$$

and use Proposition 6 to define a valid set of cuts for each $\theta_j$ separately. Indeed,

$$\theta_{rj} = p_j P\left( \sum_{i \in I(j)} \xi_{ij} < d_j \right) \tag{29}$$

and

$$\lambda_{rj}(1) = p_j \min_{i \notin I(j)} P\left[\left(\sum_{i \in I(j)} \xi_{ij}\right) + \xi_{tj} < d_j\right]. \tag{30}$$

Note that in practice, finding $t$ can be relatively easy. For instance, if all random variables are Poisson and independent, $t$ is given simply by the random variable $\xi_{it}(t \notin I(j))$ with the largest parameter value. Similarly, we may use Corollary 4 separately for each component $j$. For a given $x^k$,

$$R_j\left(x^k\right) = p_j E_\xi \left(d_j - \sum_{i \in I(j)} \xi_{ij} {x_{ij}}^k\right)^+ / d_j.$$

Using again the notation $I(j) = \left\{i : x_{ij}^k = 1\right\}$, the $i^{th}$ component of $\partial R_j\left(x^k\right)$ is equal to

$$(-p_j/d_j)E\left[\xi_{ij} P\left(\sum_{i \in I(j)} \xi_{ij} < d_j\right)\right] \quad \text{if} \quad i \in I(j),$$

and to

$$(-p_j/d_j)E_{D_j}(\xi_{ij}) \text{ if } i \notin I(j),$$

with

$$D_j = \left\{\xi_{ij} : \sum_{i \in I(j)} \xi_{ij} \leq d_j\right\}.$$

To illustrate, consider the following numerical example. Let $n = 4, m = 6, f_i = 10$ for all $i$ and $p_j = 40$ for all $j$. The $c_{ij}$ coefficients are given by

| $i \setminus j$ | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| 1 | 10 | 12 | 8 | 6 | 5 | 14 |
| 2 | 8 | 5 | 10 | 15 | 9 | 12 |
| 3 | 7 | 14 | 4 | 11 | 15 | 8 |
| 4 | 5 | 8 | 12 | 10 | 10 | 10 |

The $\xi_{ij}$ are independent Poisson variables with parameters

| $i \setminus j$ | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| 1 | 4 | 4 | 5 | 3 | 3 | 8 |
| 2 | 5 | 2 | 4 | 8 | 5 | 6 |
| 3 | 2 | 8 | 3 | 4 | 7 | 5 |
| 4 | 3 | 5 | 6 | 4 | 6 | 5 |

and, finally, the demands $d_j$ are given by

| $j$ | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| $d_j$ | 8 | 4 | 6 | 3 | 5 | 8 |

We illustrate the generation of the cuts for $j = 1$. First, a lower bound is obtained by letting $x_{i1} = 1$ for all $i$. This gives $L_1 = 1.265$. Assume a starting solution $x_{ij} = 0$ for all $i, j$. Cut (18) becomes $\theta_1 \geq 40 - 19.368\,(x_{11} + x_{21} + x_{31} + x_{41})$, with the coefficient $a = 19.368$ obtained from $(\theta_{r1} - L_1)/2$, while the continuous cut is $\theta_1 \geq 40 - 20\,x_{11} - 25x_{21} - 10x_{31} - 15x_{41}$. The next iterate point includes $x_{11} = 1, x_{21} = 0, x_{31} = 0, x_{41} = 1$. Cut (18) becomes $\theta_1 \geq -16.788 + 20.368\,(x_{11} - x_{21} - x_{31} + x_{41})$, with the coefficient $a$ now obtained from $\theta_{r1} - \lambda_{r1}(1)$, while the continuous cut (16) is $\theta_1 \geq 29.164 - 11.974\,x_{11} - 14.968\,x_{21} - 5.987\,x_{31} - 8.981\,x_{41}$. Cut (18) is stronger than (16) at the current iterate point, with value 23.948 instead of 8.309. Also, as the coefficient $a$ comes from $\theta_{r1} - \lambda_{r1}(1)$ and $\lambda_{r1}(1)$ is obtained when $x_{21}$ becomes 1, inequality (18) gives an exact bound on the solution $x_{11} = 1, x_{21} = 1, x_{31} = 0, x_{41} = 1$. It provides a non-trivial but non-binding bound for other cases, such as $x_{11} = 0, x_{21} = x_{31} = x_{41} = 1$. On the other hand, (18) provides a non-trivial (but non-binding) bound for some cases such as $x_{11} = 0, x_{21} = 1, x_{31} = 1, x_{41} = 0$, whereas (16) does not.

Cuts (16) and (18) were added each time the amount of violation exceeded .1. The number of iterate points is dependent on the branch-and-bound strategies. For this example, the largest

number of iterate points was 21. In that case, the mean number of cuts per component was 6.833 cuts of type (16) and 2.5 cuts (18). As extreme cases, 10 integrality cuts were imposed for output 1 and only 4 for output 2, while 4 continuous cuts were imposed for output 3 and only 1 for output 5.

The algorithm was simulated by adding cuts each time a new iterate point was found, then restarting the branch-and-bound. An estimate of the time needed is obtained by summing up the incremental times between consecutive iterate points. This yielded a total of 44 minutes on a 386 PC, 16 Mhz with coprocessor.

The optimal solution is $x_{11} = x_{13} = x_{15} = x_{16} = x_{21} = x_{22} = x_{24} = x_{41} = x_{42} = x_{43} = x_{45} = x_{46} = 1$ and all other $x_{ij}$'s are 0 with first-stage cost 140, and penalty 13.261, for a total of 153.261. It strongly differs from the solution of the deterministic problem where outputs equal expected values: $x_{11} = x_{12} = x_{13} = x_{14} = x_{16} = x_{21} = x_{23} = x_{25} = 0$ with first-stage cost 97 and no penalty. The reason is that in the stochastic case, even if the expected output exceeds demand, there still exists a probability that demand will not be met.

**Acknowledgements**

This work was partially supported by the Canadian Natural Sciences and Engineering Research Council (grant OGP0039682) and by the Centre Interuniversitaire dans les Sciences doctorales du Management, Belgium. Thanks are due to an anonymous referee for his valuable comments.

## References

J.F. BENDERS, "Partitioning Procedures for Solving Mixed-Variables Programming Problems", *Numerische Mathematik* 4 (1962), 238–252.

D. BIENSTOCK and J.F. SHAPIRO, "Optimizing Resource Acquisition Decisions by Stochastic Programming", *Management Science* 34 (1988), 215–229.

J.R. BIRGE and F.V. LOUVEAUX, "A Multicut Algorithm for Two-Stage Stochastic Linear Programs", *European Journal of Operational Research* 34 (1988), 384–392.

J.R. BIRGE and R. J-B WETS, "Designing Approximation Schemes for Stochastic Optimization Problems, in Particular for Stochastic Programs with Recourse", *Mathematical Programming Study* 27 (1986), 54–102.

M. DROR and P. TRUDEAU, "Stochastic Vehicle Routing with Modified Savings Algorithm", *European Journal of Operations Research*, 23, (1986), 228–235.

Y. ERMOLIEV and R. J-B WETS (eds.), "Numerical Techniques for Stochastic Optimization", Springer-Verlag, Berlin (1988).

P. JAILLET, "A Priori Solution of a Travelling Salesman Problem in which a Random Subset of the Customers are Visited", *Operations Research* 36 (1988), 929–936.

G. LAPORTE and F.V. LOUVEAUX, "Formulation and Bounds for the Stochastic Capacitated Vehicle Routing Problem with Uncertain Supplies", in J.J. Gabszewicz, J.F. Richard and L.A. Wolsey (eds.), *Economic Decision Making: Games, Econometrics and Optimization*, North-Holland (1990), 443–455.

G. LAPORTE, F.V. LOUVEAUX and H. MERCURE, "The Vehicle Routing Problem with Stochastic Travel Times", *Transportation Science* 26 (1992), 161–170.

G. LAPORTE, F.V. LOUVEAUX and L. VAN HAMME, "Exact Solution of a Stochastic Location Problem by an Integer L-shaped Algorithm", C.R.T. #716, Université de Montréal (1992).

F.V. LOUVEAUX, "Multistage Stochastic Programs with Block-Separable Recourse", *Mathematical Programming Study* 28 (1986), 48–62.

H.N. PSARAFTIS, "On the Practical Importance of Asymptotic Optimality in Certain Heuristic Algorithms", *Networks*, 14, (1984), 587–596.

STOUGIE, L., "Design and Analysis of Algorithms for Stochastic Integer Programming", Ph.D. Thesis, Centre for Mathematics and Computer Science, Amsterdam (1985).

R. VAN SLYKE and R.J-B WETS, "L-shaped Linear Programs with Applications to Optimal Control and Stochastic Programming", *SIAM Journal on Applied Mathematics* 17 (1969), 638–663.

R. J-B WETS, "Stochastic Programming: Solution Techniques and Approximation Schemes", in A. Bachem, M. Grötschel and B. Korte (eds.), Mathematical Programming: The State of the Art, Springer-Verlag, Berlin (1983), 566–603.

R.D. WOLLMER, "Two-stage Linear Programming under Uncertainty with 0–1 Integer First-Stage Variables", *Mathematical Programming* 19 (1980), 279–288.